

3-1-1995

Generalized Sylvester theorems for periodic applications in matrix optics

Lee W. Casperson
Portland State University

Anthony A. Tovar

Let us know how access to this document benefits you.

Follow this and additional works at: http://pdxscholar.library.pdx.edu/ece_fac



Part of the [Electrical and Computer Engineering Commons](#)

Citation Details

Anthony A. Tovar and W. Casperson, "Generalized Sylvester theorems for periodic applications in matrix optics," J. Opt. Soc. Am. A 12, 578-590 (1995).

This Article is brought to you for free and open access. It has been accepted for inclusion in Electrical and Computer Engineering Faculty Publications and Presentations by an authorized administrator of PDXScholar. For more information, please contact pdxscholar@pdx.edu.

Generalized Sylvester theorems for periodic applications in matrix optics

Anthony A. Tovar and Lee W. Casperson

Department of Electrical Engineering, Portland State University, Portland, Oregon 97207-0751

Received June 9, 1994; revised manuscript received September 19, 1994; accepted September 20, 1994

Sylvester's theorem is often applied to problems involving light propagation through periodic optical systems represented by unimodular 2×2 transfer matrices. We extend this theorem to apply to broader classes of optics-related matrices. These matrices may be 2×2 or take on an important augmented 3×3 form. The results, which are summarized in tabular form, are useful for the analysis and the synthesis of a variety of optical systems, such as those that contain periodic distributed-feedback lasers, lossy birefringent filters, periodic pulse compressors, and misaligned lenses and mirrors. The results are also applicable to other types of system such as periodic electric circuits with intracavity independent sources, high-energy particle accelerators, and periodic computer graphics manipulations that may include object translation. As an example, we use the 3×3 form of Sylvester's theorem to examine Gaussian beam propagation in a misaligned resonator.

1. INTRODUCTION

Ubiquitous 2×2 transfer-matrix methods are commonly used in the study of a wide variety of problems in optics^{1,2} and other areas of engineering and physics. With these methods system analysis involves only 2×2 matrix multiplication. When the optical system under consideration has properties that vary periodically, the system matrix must be multiplied by itself several times. The mathematical formula governing this procedure is known as Sylvester's theorem. Typically, the somewhat obscure general form of Sylvester's theorem is reported only in mathematical texts, or else one encounters the simplest explicit special case for unimodular 2×2 matrices. However, it has long been desired to obtain a practical form of Sylvester's theorem for six-element 3×3 matrices.³ One purpose of this paper is to develop Sylvester's theorem for such matrices having arbitrary determinants.

Periodic systems occur naturally as in, for example, homogeneous crystals that contain periodic crystalline planes. Similarly, periodic sequences of lenses or apertures may be used for optical waveguiding.^{4,5} In the limit of a small period these lens waveguides have the same propagation characteristics as those of inhomogeneous lenslike media. Conventional lasers, Fabry-Perot interferometers, optical delay lines,⁶ and multipass resonators⁷ represent important classes of periodic optical systems that may have the same mode structure as that of these lens waveguides. Periodically perturbed optical fibers are considered to be a possible consequence of defective manufacturing techniques or distortions in multifiber cables.⁸ Acoustic waves are sometimes used to generate a periodic refractive profile in acousto-optic media. An important class of particle accelerators is periodic.^{9,10} In addition to its role in applications with these systems, Sylvester's theorem may also be used in the analysis and the design of fan and folded Solc filters, distributed-feedback waveguides and lasers, twisted

nematic liquid crystals, Bragg filters, and surface wave devices.¹¹

Sylvester's theorem is usually reported only for the special case of unimodular matrix theories. However, many matrix theories are unimodular only when the optical system is subject to some restrictive constraints. For example, Jones calculus¹² is unimodular only when the optical system is lossless and when absolute phase is ignored. This situation is certainly invalid when the optical system contains polarizing elements or when polarization-dependent reflection and refraction for non-normal incidence are accounted for.¹³ Only the assumption of losslessness permits other matrix methods, such as those involving characteristic matrices for stratified media,¹⁴ transfer matrices for distributed-feedback structures,¹⁵ and transfer matrices for fiber ring resonators¹⁶ to be unimodular as well.

For every 2×2 transfer-matrix method there is a corresponding augmented 3×3 matrix method.² The form of the 3×3 matrix of interest here, however, contains only six independent elements. In both paraxial ray matrix theory and Gaussian beam theory the 3×3 matrix method permits the designer to trace paraxial light rays and Gaussian beams through misaligned optical systems.³ This 3×3 formalism may also be applied to, for example, the design of pulse compressors,¹⁷ dispersive laser cavities,¹⁸ and particle accelerators.¹⁹ Similarly, 3×3 matrix methods are necessary for the study of two-port electrical circuits that contain intranetwork independent voltage and current sources and for computer graphics manipulations in which object translation is required.²⁰

Many of the generalized Sylvester theorems of interest here are derivable from a still more general form that we will refer to as Sylvester's matrix polynomial theorem. As an example, the conventional 2×2 unimodular Sylvester theorem is derived from Sylvester's matrix polynomial theorem in Section 2. Procedures for applying Sylvester's theorem and several special cases

of the theorem are also identified there. In Section 3 Sylvester's theorem is generalized to apply to a variety of periodic systems that are represented by matrices that may be nonunimodular, zero determinant, 3×3 , or any combination thereof. In Section 4 the use of the generalized Sylvester theorems is demonstrated with an example.

2. SYLVESTER'S MATRIX POLYNOMIAL THEOREM

In transfer-matrix applications, system analysis involves simple matrix multiplication. When the system under consideration is periodic, the system matrix must be multiplied by itself several times, and the formula for the s th power of an $n \times n$ matrix is of interest. This formula, named after the 19th-century mathematician James Joseph Sylvester (1814–1897), is Sylvester's matrix polynomial theorem, reported in 1882 (Ref. 21) and sometimes referred to as Sylvester's theorem,²² Sylvester's formula,²³ or the Lagrange–Sylvester formula.²⁴ Sylvester, a pioneer in linear algebra, is also responsible for the term “matrix,” his usage of which began in 1850. There are differing opinions on the name to be given to the 2×2 special case of Sylvester's matrix polynomial theorem. The reason may stem in

evaluate Sylvester's theorem in the most general case is also discussed. This process, though it is straightforward in principle, is sometimes difficult in practice. Thus several special cases of Sylvester's theorem are identified in this section.

Suppose that T is an $n \times n$ matrix, I is the $n \times n$ identity matrix, and λ is a scalar. A characteristic equation may be defined in terms of a determinant:

$$p(\lambda) = |T - \lambda I| = 0. \quad (1)$$

The scalar roots of this characteristic equation, $\lambda_1, \lambda_2, \dots, \lambda_n$, are the eigenvalues of the matrix T . If $p(\lambda)$ has no multiple roots, the s th power of the matrix T is

$$T^s = \sum_{j=1}^n \left(\lambda_j^s \prod_{i \neq j} \frac{\lambda_i I - T}{\lambda_i - \lambda_j} \right). \quad (2)$$

This is Sylvester's matrix polynomial theorem. When the eigenvalue equation (1) has nondistinct roots, there is a more general confluent form of the theorem.³³

A. Sylvester's Matrix Polynomial

Theorem for 2×2 Matrices

To illustrate the usage of Eq. (2), we consider the simplest nontrivial matrix exponent. For this 2×2 matrix case $n = 2$, and Eq. (2) reduces to

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^s = \sum_{j=1}^2 \left(\lambda_j^s \prod_{i \neq j} \frac{1}{\lambda_i - \lambda_j} \begin{bmatrix} \lambda_i - A & -B \\ -C & \lambda_i - D \end{bmatrix} \right), \quad (3)$$

$$= \frac{\lambda_1^s}{\lambda_2 - \lambda_1} \begin{bmatrix} \lambda_2 - A & -B \\ -C & \lambda_2 - D \end{bmatrix} + \frac{\lambda_2^s}{\lambda_1 - \lambda_2} \begin{bmatrix} \lambda_1 - A & -B \\ -C & \lambda_1 - D \end{bmatrix}, \quad (4)$$

$$= \begin{bmatrix} A \frac{\lambda_2^s - \lambda_1^s}{\lambda_2 - \lambda_1} + \lambda_1 \lambda_2 \frac{\lambda_1^{s-1} - \lambda_2^{s-1}}{\lambda_2 - \lambda_1} & B \frac{\lambda_2^s - \lambda_1^s}{\lambda_2 - \lambda_1} \\ C \frac{\lambda_2^s - \lambda_1^s}{\lambda_2 - \lambda_1} & D \frac{\lambda_2^s - \lambda_1^s}{\lambda_2 - \lambda_1} + \lambda_1 \lambda_2 \frac{\lambda_1^{s-1} - \lambda_2^{s-1}}{\lambda_2 - \lambda_1} \end{bmatrix}. \quad (5)$$

part from the fact that this important case of the theorem was reported in 1858 by Sylvester's close friend Arthur Cayley (1821–1895). In fact, Cayley, the creator of modern matrix theory, had written the his original paper on matrices.²⁵ In a standard optics 2×2 theorem in text Born and Wolf report the theorem but do not give it a name,²⁶ and others have done the same.^{27,28} Perhaps because the 2×2 theorem can be written in terms of Chebyshev polynomials, it has recently been called Chebyshev's identity,²⁹ and others have followed in this usage.^{11,15} The most common and more appropriate name used, however, is Sylvester's theorem,^{1–3,30–32} and this convention is followed here. Specifically, we will mean by Sylvester's theorem the 2×2 unimodular (unit-determinant) special case of the more general Sylvester matrix polynomial theorem. (It may be noted that Sylvester is responsible for many other theorems, several of which are sometimes also referred to non-descriptively as Sylvester's theorem.)

Many of the results of this study are derivable directly from Sylvester's matrix polynomial theorem. While this general theorem is known in the mathematics literature, its relevance in optics is not so well known. As an example of its use, Sylvester's theorem (2×2 unimodular case) is derived from it in this section. The procedure to

Given the two eigenvalues of the matrix T , Eq. (5) is the formula for T^s . The eigenvalues, as we mentioned above, can be obtained from the characteristic equation (1), which for a 2×2 matrix is

$$p(\lambda) = (A - \lambda)(D - \lambda) - BC = 0. \quad (6)$$

Equation (6) may be rewritten in standard quadratic equation form:

$$\lambda^2 - (A + D)\lambda + (AD - BC) = 0. \quad (7)$$

There are several ways to proceed. It is clear that the trace $A + D$ and the determinant $AD - BC$ are important quantities, and the eigenvalues depend directly on these quantities. For this derivation only, the frequently occurring special case in which the determinant is unity is considered. The trace is permitted to be arbitrary. If the definition

$$\cos \theta \equiv \frac{A + D}{2} \quad (8)$$

is introduced, then the eigenvalues from the quadratic characteristic equation (7) are

$$\lambda_{1,2} = \exp(\pm i\theta), \quad (9)$$

and Eq. (5) reduces to

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^s = \frac{1}{\sin \theta} \times \begin{bmatrix} A \sin(s\theta) - \sin[(s-1)\theta] & B \sin(s\theta) \\ C \sin(s\theta) & D \sin(s\theta) - \sin[(s-1)\theta] \end{bmatrix}, \quad (10)$$

which is the standard form of Sylvester's theorem.

B. Alternative Forms and Evaluation of Sylvester's Theorem

To emphasize the polynomial nature of the solution, one sometimes writes Eq. (10) in terms of Chebyshev polynomials of the second kind²⁶:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^s = \begin{bmatrix} AU_{s-1}(x) - U_{s-2}(x) & BU_{s-1}(x) \\ CU_{s-1}(x) & DU_{s-1}(x) - U_{s-2}(x) \end{bmatrix}, \quad (11)$$

where

$$x = 1/2(A + D), \quad (12)$$

$$U_s(x) = \frac{\sin[(s+1)\cos^{-1}(x)]}{(1-x^2)^{1/2}}. \quad (13)$$

The first several Chebyshev polynomials are

$$\begin{aligned} U_0(x) &= 1, \\ U_1(x) &= 2x, \\ U_2(x) &= 4x^2 - 1, \\ U_3(x) &= 8x^3 - 4x, \\ U_4(x) &= 16x^4 - 12x^2 + 1, \\ U_5(x) &= 32x^5 - 32x^3 + 6x, \\ U_6(x) &= 64x^6 - 80x^4 + 24x^2 - 1. \end{aligned} \quad (14)$$

Additional polynomials may be obtained from the recursion relation³⁴

$$U_{n+1} - 2xU_n + U_{n-1} = 0. \quad (15)$$

It is clear that Eqs. (11)–(13) are equivalent to Eqs. (8) and (10).

The unimodular 2×2 form of Sylvester's theorem has

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^s = \frac{1}{\sinh \phi} \begin{bmatrix} A \sinh(s\phi) - \sinh[(s-1)\phi] & B \sinh(s\phi) \\ C \sinh(s\phi) & D \sinh(s\phi) - \sinh[(s-1)\phi] \end{bmatrix}, \quad (20)$$

been derived above from Sylvester's matrix polynomial theorem. Alternatively, one may derive it from first principles, and for completeness this derivation is given in Appendix A below. A similar derivation is also given in Ref. 1. Conceptually simpler is the inductive proof, which is given in Appendix B.³² We also use these two appendixes to explore the range of validity of Sylvester's theorem. For example, these appendixes suggest that Sylvester's theorem may also apply to roots of matrices. In Appendix C roots are specifically considered, and it is

shown that Sylvester's theorem is also valid for these matrix roots.^{22,25} The fact that Sylvester's theorem applies to integer powers and roots of matrices suggests that it applies to arbitrary rational powers. The proof of this is given in Appendix D. The proofs in these appendixes are crucial for a rigorous understanding of Sylvester's theorem and its applications to matrix optics. With these proofs this section contains a comprehensive summary of the properties of this important theorem.

In general, each of the $ABCD$ matrix elements in Sylvester's theorem may be complex. We begin the procedure for evaluating matrix equation (10) by calculating θ . One may determine this complex angle explicitly by combining the Euler relation with Eq. (8) to yield

$$\theta = -i \ln \left\{ \frac{A+D}{2} + \left[\left(\frac{A+D}{2} \right)^2 - 1 \right]^{1/2} \right\}. \quad (16)$$

The complex square root in this formula may be evaluated with the use of

$$(a + ib)^{1/2} = \left[\frac{(a^2 + b^2)^{1/2} + a}{2} \right]^{1/2} + i(\operatorname{sgn} b) \left[\frac{(a^2 + b^2)^{1/2} - a}{2} \right]^{1/2}, \quad (17)$$

where the signum function $\operatorname{sgn} b \equiv b/|b|$ has been used. We may separate the complex natural logarithm into real and imaginary parts, using the relationship

$$\ln(a + ib) = \ln[(a^2 + b^2)^{1/2}] + i[\tan^{-1}(b/a)]. \quad (18)$$

Finally, one may evaluate the elements of matrix equation (10) by noting that

$$\sin(a + ib) = \sin a \cosh b + i \cos a \sinh b. \quad (19)$$

This procedure for evaluating Eq. (10) simplifies if $(A + D)/2$ is real. However, if $(A + D)/2$ is greater than unity in magnitude, then, from Eq. (16), θ is purely imaginary. In this case it is convenient to rewrite Sylvester's theorem in another form. If $\theta \equiv i\phi$, then Eq. (10) may be written as

where

$$\cosh \phi = \frac{A+D}{2}. \quad (21)$$

C. Special Cases of Sylvester's Theorem

Because of the somewhat complicated form of Sylvester's theorem, it is useful to consider special cases of Eq. (10) in

which the matrix elements take on specific forms. One such special case exists when $A = D \equiv \cos \theta$ and $\chi \equiv (-B/C)^{1/2}$. Here Sylvester's theorem reduces to

$$\begin{bmatrix} \cos \theta & \chi \sin \theta \\ -\chi^{-1} \sin \theta & \cos \theta \end{bmatrix}^s = \begin{bmatrix} \cos(s\theta) & \chi \sin(s\theta) \\ -\chi^{-1} \sin(s\theta) & \cos(s\theta) \end{bmatrix}. \quad (22)$$

In the limit as θ approaches zero, this equation has two forms of interest:

$$\begin{bmatrix} 1 & \chi_1 \\ 0 & 1 \end{bmatrix}^s = \begin{bmatrix} 1 & s\chi_1 \\ 0 & 1 \end{bmatrix}, \quad (23)$$

$$\begin{bmatrix} 1 & 0 \\ \chi_2 & 1 \end{bmatrix}^s = \begin{bmatrix} 1 & 0 \\ s\chi_2 & 1 \end{bmatrix}. \quad (24)$$

For nonunimodular matrices an important special case of Sylvester's matrix polynomial theorem is

$$\begin{bmatrix} \chi_1 & 0 \\ 0 & \chi_2 \end{bmatrix}^s = \begin{bmatrix} \chi_1^s & 0 \\ 0 & \chi_2^s \end{bmatrix}. \quad (25)$$

Although this result is almost obvious, it may be derived by solution of eigenvalue equation (7) and with the use of Eq. (5). The eigenvalues in this case are A and D .

If s is constrained to be an integer, then one may use Sylvester's matrix polynomial theorem to evaluate the interesting off-diagonal matrix form

$$\begin{bmatrix} 0 & \chi_1 \\ \chi_2 & 0 \end{bmatrix}^{2s} = (\chi_1 \chi_2)^s \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad (26a)$$

$$\begin{bmatrix} 0 & \chi_1 \\ \chi_2 & 0 \end{bmatrix}^{2s+1} = (\chi_1 \chi_2)^s \begin{bmatrix} 0 & \chi_1 \\ \chi_2 & 0 \end{bmatrix}, \quad (26b)$$

These results may also be found from Eqs. (5) and (7). The eigenvalues are $\pm(BC)^{1/2}$. When $\chi_1 \chi_2 = -1$, these results follow from Eq. (22) with $\theta = \pi/2$. The matrix equations (26) are of special interest in the analysis of confocal resonators. In this case this off-diagonal matrix represents the transformation of a Gaussian beam after it propagates from the center of the resonator back to the center of the resonator.

In the synthesis of both multipass and periodic systems, factorizations and roots of matrices are often of interest. For example, it is sometimes desired that an optical system be transparent in the sense that the system matrix is made to be the identity matrix. Some factorizations of the identity matrix include²

$$\begin{aligned} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} &= \left[\begin{bmatrix} 1 & 0 \\ -3/l & 1 \end{bmatrix} \begin{bmatrix} 1 & l \\ 0 & 1 \end{bmatrix} \right]^3 = \left[\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right]^2 \\ &= \left[\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \right]^2 = \begin{bmatrix} 0 & \gamma \\ \gamma^{-1} & 0 \end{bmatrix}^2 \\ &= - \begin{bmatrix} 0 & \gamma \\ -\gamma^{-1} & 0 \end{bmatrix}^2, \end{aligned} \quad (27)$$

where γ is permitted to be complex. The identity matrix taken to any positive integer power is the identity matrix, and this result may be obtained from either Eq. (23) or

(24). However, it is clear from Eq. (27) that the identity matrix has an infinite number of distinct square roots (and cube roots too). These do not follow from Eq. (23) or (24), since it was assumed that the eigenvalues in Eq. (9) were distinct. The difficulty may be readily seen from Appendix A below. Since the eigenvalues are equal, the eigenvalue matrix Λ commutes with the matrix M . In this case the substitution $T = M^{-1}\Lambda M$ reduces to $T = \Lambda$, which is not useful, and an alternative technique must be employed. In particular, roots must be examined specifically,

$$T^s = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad (28)$$

and Sylvester's theorem is well suited to this problem. We may use the unimodular form of Sylvester's theorem to reduce this equation to four scalar equations with four unknowns:

$$\frac{A \sin(s\theta) - \sin[(s-1)\theta]}{\sin \theta} = 1, \quad (29)$$

$$\frac{C \sin(s\theta)}{\sin \theta} = 0, \quad (30)$$

$$\frac{B \sin(s\theta)}{\sin \theta} = 0, \quad (31)$$

$$\frac{D \sin(s\theta) - \sin[(s-1)\theta]}{\sin \theta} = 1, \quad (32)$$

where θ is defined in Eq. (16). The trivial solutions are $A = D = \pm 1$ and $B = C = 0$. Alternatively, Eqs. (30) and (31) are satisfied if

$$\sin(s\theta) = 0, \quad (33)$$

$$\sin \theta \neq 0. \quad (34)$$

After the use of a trigonometric identity it may be seen that Eqs. (29) and (32) are solved when

$$\cos(s\theta) = 1. \quad (35)$$

Equations (33) and (35) are satisfied when

$$s\theta = 2k\pi, \quad (36)$$

where k may be any integer. To avoid duplicating solutions, we impose the restriction⁷

$$0 \leq k \leq s/2. \quad (37)$$

We may combine Eq. (34) with Eq. (36) to obtain a restriction on s , and it follows that $s \neq 2$. Thus the identity matrix has no nontrivial unimodular square roots.

However, there are nontrivial nonunimodular square roots, those with a determinant of -1 . As we can show from the nonunimodular Sylvester theorem in Section 3 below, the third criterion [Eq. (35)] becomes

$$\cos(s\theta) = (AD - BC)^{-s/2}. \quad (38)$$

If $s = 2$ and the determinant is -1 , then the three criteria

[Eqs. (33), (34), and (38)] are satisfied when $\theta = \pi/2$. The nonunimodular relationship

$$\cos \theta = 1/2(A + D)(AD - BC)^{-1/2} \quad (39)$$

may be used, and it follows that, for this value of θ , $D = -A$. Since the determinant is -1 , the result may be written as

$$\sqrt{I} = \begin{bmatrix} \cos \phi & \gamma \sin \phi \\ \gamma^{-1} \sin \phi & -\cos \phi \end{bmatrix}. \quad (40)$$

In the special case $\phi = \pi/2$ Eq. (40) reduces to the fourth equality in Eq. (27). As an additional result, if a matrix T is a known sth root of the identity matrix, then $\sqrt{I} T \sqrt{I}$ is also an sth root of the identity matrix. Since a nonunimodular Sylvester's theorem has been used, these results provide an additional motivating factor for the derivation of generalized Sylvester's theorems. These generalized theorems are considered in Section 3.

3. GENERALIZED SYLVESTER THEOREMS

It has long been desired to obtain Sylvester's theorem for six-element 3×3 matrices.³ The purpose of this section is to derive nonunimodular 2×2 and 3×3 forms of Sylvester's theorem. The methodology used is discussed, and it is straightforward to extend the results here to other higher-order matrices.

A. Nonunimodular 2×2 Sylvester Theorems

There is more than one method that we may use to find the nonunimodular form of Sylvester's theorem. From Sylvester's matrix polynomial theorem it is clear that the desired matrix may be obtained from Eq. (5) with the eigenvalues derived from the general solution to the characteristic equation (7). One may also derive it directly in the same manner as is done in Appendix A below. Alternatively, if the determinant $\tau = AD - BC$ is nonzero, it may be factored out, and

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^s = \tau^{s/2} \begin{bmatrix} A\tau^{-1/2} & B\tau^{-1/2} \\ C\tau^{-1/2} & D\tau^{-1/2} \end{bmatrix}^s. \quad (41)$$

Now, Sylvester's theorem [Eq. (10)] may be applied to the matrix on the right-hand side of Eq. (41), since it is now unimodular:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^s = \frac{\tau^{(s-1)/2}}{\sin \theta'} \begin{bmatrix} A \sin(s\theta') - \sqrt{\tau} \sin[(s-1)\theta'] & B \sin(s\theta') \\ C \sin(s\theta') & D \sin(s\theta') - \sqrt{\tau} \sin[(s-1)\theta'] \end{bmatrix}. \quad (42)$$

The angle θ' is defined by the relationship

$$\cos \theta' = 1/2(A + D)\tau^{-1/2}. \quad (43)$$

Birefringent optical systems often include the use of polarizers. In filter design, for example, one uses polarizers to discard unwanted frequency components. In the Jones calculus, polarizers are represented by zero-determinant matrices. Since the determinant of a product is the product of the determinants, it follows that any optical system that includes polarizers would be represented by a zero-determinant matrix. Thus it is important to derive a zero-determinant form of Sylvester's theorem. From

Eq. (7) it follows that the eigenvalues are $A + D$ and zero, and from Eq. (5) it is seen that

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^s = (A + D)^{s-1} \begin{bmatrix} A & B \\ C & D \end{bmatrix}. \quad (44)$$

The results of this subsection are summarized in Table 1. Because of the generality of the nonunimodular form of Sylvester's theorem, it is appropriate to examine special cases. Thus several specific matrix operations that have different exponents are also identified in Table 1. Other matrix operations that may be derived directly from Sylvester's theorem may also be derived from the special-case matrices. For example, the square-root matrix in Table 1 may be applied to itself to yield the fourth-root matrix operation:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{1/4} = \frac{1^{1/4}}{\sqrt{\delta}(2\tau^{1/4} + \sqrt{\delta})^{1/2}} \times \begin{bmatrix} A + \sqrt{\tau} + \tau^{1/4}\sqrt{\delta} & B \\ C & D + \sqrt{\tau} + \tau^{1/4}\sqrt{\delta} \end{bmatrix}, \quad (45)$$

where

$$\delta \equiv A + D + 2\sqrt{\tau}. \quad (46)$$

The determinant of the original matrix is $\tau = AD - BC$. Thus the determinant of the matrix in Eq. (45) is $\tau^{1/4}$. It is important to note that, according to Sylvester's theorem, a matrix taken to the -1 power is equivalent to the matrix inverse, and a matrix taken to the zero power is the identity matrix, as we would expect.

B. Nonunimodular 3×3 Sylvester Theorem

The 2×2 Sylvester theorems given above apply to a wide variety of problems in optics and physics in general. However, as we mentioned in Section 1, there are important cases in which a 3×3 matrix is needed. For example, with the 3×3 theory one may trace light rays and Gaussian beams through misaligned optical systems. Therefore there exists a need for a 3×3 version of Sylvester's theorem for the three cases given above: nonunimodular, unimodular, and zero-determinant ma-

trices. The general nine-element 3×3 form of Sylvester's theorem is not necessary, and we are concerned with 3×3 matrices that have the form

$$\begin{bmatrix} X_s \\ Y_s \\ 1 \end{bmatrix} = \begin{bmatrix} A & B & E \\ C & D & F \\ 0 & 0 & 1 \end{bmatrix}^s \begin{bmatrix} X_0 \\ Y_0 \\ 1 \end{bmatrix} = \begin{bmatrix} A_s & B_s & E_s \\ C_s & D_s & F_s \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_0 \\ Y_0 \\ 1 \end{bmatrix}. \quad (47)$$

Regardless of the determinant, the A_s , B_s , C_s , and D_s terms of the sth power of the 3×3 matrix equation (47) are the same as their 2×2 equivalents [Eq. (42)]. Thus

Table 1. Generalized 2×2 Sylvester Theorems

Description	Operation	Matrix	
		$\tau \equiv AD - BC$	$\cos \theta \equiv 1/2(A + D)\tau^{-1/2}$
Sylvester's theorem ($\tau \neq 0$)	$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^s$	$\frac{\tau^{(s-1)/2}}{\sin \theta} \begin{bmatrix} A \sin(s\theta) - \sqrt{\tau} \sin[(s-1)\theta] & B \sin(s\theta) \\ C \sin(s\theta) & D \sin(s\theta) - \sqrt{\tau} \sin[(s-1)\theta] \end{bmatrix}$	
Sylvester's theorem ($\tau = 1$)	$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^s$	$\frac{1}{\sin \theta} \begin{bmatrix} A \sin(s\theta) - \sin[(s-1)\theta] & B \sin(s\theta) \\ C \sin(s\theta) & D \sin(s\theta) - \sin[(s-1)\theta] \end{bmatrix}$	
Sylvester's theorem ($\tau = 0$)	$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^s$	$(A + D)^{s-1} \begin{bmatrix} A & B \\ C & D \end{bmatrix}$	
Squared matrix	$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^2$	$\begin{bmatrix} A(A + D) - \tau & B(A + D) \\ C(A + D) & D(A + D) - \tau \end{bmatrix}$	
Unit matrix	$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^1$	$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$	
Square-root matrix	$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{1/2}$	$\frac{\pm 1}{(A + D + 2\sqrt{\tau})^{1/2}} \begin{bmatrix} A + \sqrt{\tau} & B \\ C & D + \sqrt{\tau} \end{bmatrix}$	
Identity matrix	$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^0$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	
Inverse square-root matrix	$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1/2}$	$\frac{\pm 1}{\sqrt{\tau}(A + D + 2\sqrt{\tau})^{1/2}} \begin{bmatrix} D + \sqrt{\tau} & -B \\ -C & A + \sqrt{\tau} \end{bmatrix}$	
Inverse matrix	$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1}$	$\frac{1}{\tau} \begin{bmatrix} D & -B \\ -C & A \end{bmatrix}$	
Inverse squared matrix	$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-2}$	$\frac{1}{\tau^2} \begin{bmatrix} D(A + D) - \tau & -B(A + D) \\ -C(A + D) & A(A + D) - \tau \end{bmatrix}$	

it is required to find only the E_s and F_s elements. As above, one may use Sylvester's matrix polynomial theorem [Eq. (2)] directly or generalize the derivation given in Appendix A below. Alternatively, one may use the commutativity requirement $T^s T = T T^s$ to obtain two equations with two unknowns:

$$A_s E + B_s F + E_s = A E_s + B F_s + E, \quad (48a)$$

$$C_s E + D_s F + F_s = C E_s + D F_s + F. \quad (48b)$$

Solving these two equations for E_s and F_s for a unimodular matrix yields

$$\begin{bmatrix} A & B & E \\ C & D & F \\ 0 & 0 & 1 \end{bmatrix}^s = \frac{1}{\sin \theta} \begin{bmatrix} A \sin(s\theta) - \sin[(s-1)\theta] & B \sin(s\theta) & E_s' \\ C \sin(s\theta) & D \sin(s\theta) - \sin[(s-1)\theta] & F_s' \\ 0 & 0 & \sin \theta \end{bmatrix}, \quad (49)$$

where θ is defined here as in Eq. (16) and

$$E_s' \equiv \left(\frac{(A-1)\sin(s\theta) + (D-1)\{\sin[(s-1)\theta] + \sin \theta\}}{2(\cos \theta - 1)} \right) E + \left\{ \frac{\sin(s\theta) - \sin[(s-1)\theta] - \sin \theta}{2(\cos \theta - 1)} \right\} B F, \quad (50a)$$

$$F_s' \equiv \left\{ \frac{\sin(s\theta) - \sin[(s-1)\theta] - \sin \theta}{2(\cos \theta - 1)} \right\} C E + \left(\frac{(D-1)\sin(s\theta) + (A-1)\{\sin[(s-1)\theta] + \sin \theta\}}{2(\cos \theta - 1)} \right) F. \quad (50b)$$

In a manner similar to that of Subsection 3.A above, the determinant may be factored out, and we may use

Eqs. (49) and (50) to obtain Sylvester's theorem for a nonunimodular 3×3 matrix:

$$\begin{bmatrix} A & B & E \\ C & D & F \\ 0 & 0 & 1 \end{bmatrix}^s = \frac{\tau^{(s-1)/2}}{\sin \theta'} \begin{bmatrix} A \sin(s\theta') - \sqrt{\tau} \sin[(s-1)\theta'] & B \sin(s\theta') & E_s' \\ C \sin(s\theta') & D \sin(s\theta') - \sqrt{\tau} \sin[(s-1)\theta'] & F_s' \\ 0 & 0 & \sin \theta' \end{bmatrix}. \quad (51)$$

As above,

$$\cos \theta' \equiv 1/2(A + D)\tau^{-1/2}. \quad (52)$$

The elements E_s' and F_s' are

$$E_s' \equiv \left(\frac{(A\tau^{-1/2} - 1)\sin(s\theta') + (D\tau^{-1/2} - 1)\{\sin[(s-1)\theta'] + \sin \theta'\}}{2(\cos \theta' - 1)} \right) E + \left\{ \frac{\sin(s\theta') - \sin[(s-1)\theta'] - \sin \theta'}{2\sqrt{\tau}(\cos \theta' - 1)} \right\} BF, \quad (53a)$$

$$F_s' \equiv \left\{ \frac{\sin(s\theta') - \sin[(s-1)\theta'] - \sin \theta'}{2\sqrt{\tau}(\cos \theta' - 1)} \right\} CE + \left(\frac{(D\tau^{-1/2} - 1)\sin(s\theta') + (A\tau^{-1/2} - 1)\{\sin[(s-1)\theta'] + \sin \theta'\}}{2(\cos \theta' - 1)} \right) F. \quad (53b)$$

The zero-determinant form of the 3×3 Sylvester's theorem may be obtained in a similar manner. The result is

$$\begin{bmatrix} A & B & E \\ C & D & F \\ 0 & 0 & 1 \end{bmatrix}^s = (A + D)^{s-1} \begin{bmatrix} A & B & E_s' \\ C & D & F_s' \\ 0 & 0 & (A + D)^{1-s} \end{bmatrix}, \quad (54)$$

where

$$E_s' \equiv \frac{1}{(A + D)^{s-1}} \left\{ E + \left[\frac{(A + D)^{s-1} - 1}{A + D - 1} \right] (AE + BF) \right\}, \quad (55a)$$

$$F_s' \equiv \frac{1}{(A + D)^{s-1}} \left\{ F + \left[\frac{(A + D)^{s-1} - 1}{A + D - 1} \right] (CE + DF) \right\}. \quad (55b)$$

These three 3×3 forms of Sylvester's theorem, along with an inverse matrix, are summarized in Table 2.

Again, there is an interest in special cases of the theorem, and the 3×3 extension of Eq. (22) is

$$\begin{bmatrix} \cos \theta & \chi \sin \theta & E \\ -\chi^{-1} \sin \theta & \cos \theta & F \\ 0 & 0 & 1 \end{bmatrix}^s = \begin{bmatrix} \cos(s\theta) & \chi \sin(s\theta) & E_s \\ -\chi^{-1} \sin(s\theta) & \cos(s\theta) & F_s \\ 0 & 0 & 1 \end{bmatrix}, \quad (56)$$

where

$$E_s \equiv \frac{1 + \cos \theta}{2} \left[\frac{\sin(s\theta)}{\sin \theta} + \frac{1 - \cos(s\theta)}{1 + \cos \theta} \right] E + \frac{\chi \sin \theta}{2} \left[\frac{1 - \cos(s\theta)}{1 - \cos \theta} - \frac{\sin(s\theta)}{\sin \theta} \right] F, \quad (57a)$$

$$F_s \equiv \frac{1 + \cos \theta}{2} \left[\frac{\sin(s\theta)}{\sin \theta} + \frac{1 - \cos(s\theta)}{1 + \cos \theta} \right] F - \frac{\chi^{-1} \sin \theta}{2} \left[\frac{1 - \cos(s\theta)}{1 - \cos \theta} - \frac{\sin(s\theta)}{\sin \theta} \right] E. \quad (57b)$$

Two important special cases of this result occur when $\theta \rightarrow 0$:

$$\begin{bmatrix} 1 & B & E \\ 0 & 1 & F \\ 0 & 0 & 1 \end{bmatrix}^s = \begin{bmatrix} 1 & sB & sE + s(s-1)BF/2 \\ 0 & 1 & sF \\ 0 & 0 & 1 \end{bmatrix}, \quad (58)$$

$$\begin{bmatrix} 1 & 0 & E \\ C & 1 & F \\ 0 & 0 & 1 \end{bmatrix}^s = \begin{bmatrix} 1 & 0 & sE \\ sC & 1 & sF + s(s-1)CE/2 \\ 0 & 0 & 1 \end{bmatrix}. \quad (59)$$

Here the definition of χ ($\chi \sin \theta \rightarrow B$, $-\chi^{-1} \sin \theta \rightarrow C$) has been used.

4. EXAMPLE: MISALIGNED RESONATOR

Resonators are commonly used as optical delay lines and interferometers as well as in laser and lens-waveguide applications. They consist of optical elements that are

inevitably out of perfect alignment. These misalignments, whether they are accidental or intentional, are crucial to the operation of the resonator. The purpose of this section is to analyze a multipass resonator for delay line applications. This problem is well suited to paraxial ray optics and the 3×3 form of Sylvester's theorem.

For simplicity, it is assumed that the input beam to the delay line is a fundamental Gaussian beam. Thus it is of interest to consider when the position and the slope of the center of the Gaussian beam repeat themselves after s round trips. Just before the Gaussian beam returns to its original position, it encounters a hole whereby it escapes the resonator. For specificity, it is assumed that the flat left-hand mirror is tilted at some angle ϕ and that the right-hand mirror is spherical with radius R . The distance between the mirrors is d . The input plane is just to the right of the tilted mirror, and the beam is initially going to the left. For this configuration the round-trip Gaussian beam matrix for the delay line consists of only purely real elements, since aperturing effects are ignored. In only such optical systems the center of a Gaussian beam travels along paraxial light ray trajectories.³⁵ Thus one may use ray matrix techniques to trace the displacement and the slope of the center of the Gaussian beam.

The round-trip paraxial ray matrix for the system,

$$T = \begin{bmatrix} 1 & d & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2/R & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & d & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \tan 2\phi \\ 0 & 0 & 1 \end{bmatrix} \quad (60a)$$

$$= \begin{bmatrix} 1 - 2d/R & 2d(1 - d/R) & 2d(1 - d/R)\tan 2\phi \\ -2/R & 1 - 2d/R & (1 - 2d/R)\tan 2\phi \\ 0 & 0 & 1 \end{bmatrix}, \quad (60b)$$

is unimodular. After the beam traverses s round trips in the resonator, the effective ray matrix may be obtained with the aid of the newly derived Eqs. (49) and (50).

As an input condition, it is assumed that the initial ray has zero displacement and slope. In this case the output ray position after propagation through s sections of the lens waveguide is $r_s = E_s$, where E_s is defined by Eq. (47). Since Eq. (60b) is unimodular and has the property that $A = D$, then the special-case matrices (56) and (57) are of interest. In particular, it follows from Eq. (57a) that

Table 2. Generalized 3×3 Sylvester Theorems

Description	Operation	Matrix	
		$\tau \equiv AD - BC$ $\gamma \equiv A + D - 2\sqrt{\tau}$ $\cos \theta \equiv 1/2(A + D)\tau^{-1/2}$	
Sylvester's theorem ($\tau \neq 0$)	$\begin{bmatrix} A & B & E \\ C & D & F \\ 0 & 0 & 1 \end{bmatrix}^s$	$\frac{\tau^{(s-1)/2}}{\sin \theta} \begin{bmatrix} A \sin(s\theta) - \sqrt{\tau} \sin[(s-1)\theta] & B \sin(s\theta) & E_s' \\ C \sin(s\theta) & D \sin(s\theta) - \sqrt{\tau} \sin[(s-1)\theta] & F_s' \\ 0 & 0 & \tau^{(1-s)/2} \sin \theta \end{bmatrix}$	
Sylvester's theorem ($\tau = 1$)	$\begin{bmatrix} A & B & E \\ C & D & F \\ 0 & 0 & 1 \end{bmatrix}^s$	$\frac{1}{\sin \theta} \begin{bmatrix} A \sin(s\theta) - \sin[(s-1)\theta] & B \sin(s\theta) & E_s' \\ C \sin(s\theta) & D \sin(s\theta) - \sin[(s-1)\theta] & F_s' \\ 0 & 0 & \sin \theta \end{bmatrix}$	
Sylvester's theorem ($\tau = 0$)	$\begin{bmatrix} A & B & E \\ C & D & F \\ 0 & 0 & 1 \end{bmatrix}^s$	$(A + D)^{s-1} \begin{bmatrix} A & B & E_s' \\ C & D & F_s' \\ 0 & 0 & (A + D)^{1-s} \end{bmatrix}$	
Inverse matrix	$\begin{bmatrix} A & B & E \\ C & D & F \\ 0 & 0 & 1 \end{bmatrix}^{-1}$	$\frac{1}{\tau} \begin{bmatrix} D & -B & \tau^{-1}(BF - DE) \\ -C & A & \tau^{-1}(CE - AF) \\ 0 & 0 & \tau \end{bmatrix}$	
$\tau \neq 0$			
		$E_s' \equiv \left(\frac{(A\tau^{-1/2} - 1)\sin(s\theta) + (D\tau^{-1/2} - 1)\{\sin[(s-1)\theta] + \sin \theta\}}{\gamma\tau^{-1/2}} \right) E + \left\{ \frac{\sin(s\theta) - \sin[(s-1)\theta] - \sin \theta}{\gamma} \right\} BF$	
		$F_s' \equiv \left[\frac{\sin(s\theta) - \sin[(s-1)\theta] - \sin \theta}{\gamma} \right] CE + \left(\frac{(D\tau^{-1/2} - 1)\sin(s\theta) + (A\tau^{-1/2} - 1)\{\sin[(s-1)\theta] + \sin \theta\}}{\gamma\tau^{-1/2}} \right) F$	
$\tau = 1$			
		$E_s' \equiv \left(\frac{(A - 1)\sin(s\theta) + (D - 1)\{\sin[(s-1)\theta] + \sin \theta\}}{\gamma} \right) E + \left\{ \frac{\sin(s\theta) - \sin[(s-1)\theta] - \sin \theta}{\gamma} \right\} BF$	
		$F_s' \equiv \left\{ \frac{\sin(s\theta) - \sin[(s-1)\theta] - \sin \theta}{\gamma} \right\} CE + \left(\frac{(D - 1)\sin(s\theta) + (A - 1)\{\sin[(s-1)\theta] + \sin \theta\}}{\gamma} \right) F$	
$\tau = 0$			
		$E_s' \equiv (A + D)^{1-s} \left\{ E + \left[\frac{(A + D)^{s-1} - 1}{A + D - 1} \right] (AE + BF) \right\}$	
		$F_s' \equiv (A + D)^{1-s} \left\{ F + \left[\frac{(A + D)^{s-1} - 1}{A + D - 1} \right] (CE + DF) \right\}$	

$$r_s = \frac{1 + \cos \theta}{2} \left[\frac{\sin(s\theta)}{\sin \theta} + \frac{1 - \cos(s\theta)}{1 + \cos \theta} \right] E + \frac{\sin \theta}{2} \left[\frac{1 - \cos(s\theta)}{1 - \cos \theta} - \frac{\sin(s\theta)}{\sin \theta} \right] \chi F, \quad (61)$$

where $\chi = (Rd)^{1/2}(1 - d/R)^{1/2}$. With this input condition the displacement of the Gaussian beam after a single round trip is $r_1 = E$. Furthermore, it may be noted from Eq. (60b) that

$$BF = AE, \quad (62)$$

which may be written as $\chi F \sin \theta = E \cos \theta$. Thus Eq. (61) reduces to

$$\frac{2r_s}{r_1} = \frac{\sin(s\theta)}{\sin \theta} + \frac{1 - \cos(s\theta)}{1 - \cos \theta}, \quad (63)$$

which can be written as

$$r_s = r_{av} + r_{max} \sin(s\theta + \alpha), \quad (64)$$

where $r_{av} = 2^{-1}r_1(1 - \cos \theta)^{-1}$, $r_{max} = 2^{-1/2}r_1(\sin \theta)^{-1}(1 - \cos \theta)^{-1/2}$, and $\tan \alpha = -(\sin \theta)(1 - \cos \theta)^{-1}$. Except for the term r_{av} , Eq. (64) is in the same form as that for a light ray in an aligned optical system.³⁶

The 3×3 form of Sylvester's theorem has been used in demonstrating that the trajectories of the center of an initially misaligned Gaussian beam are qualitatively similar, whether they are for aligned optical systems or those with simple misalignments. However, the exact ray positions are different, and they may be obtained from Eq. (63). The qualitative result is reasonable, since the tilt of the flat mirror may be understood to have the effect of redefining the optic axis in simple two-mirror resonators such as the one considered here. For more complicated system topologies this axis redefinition may

not be possible, but such systems may still be examined with the results of the Section 3 above. This valuable method for treating misaligned elements permits the systematic treatment of multipass systems that could not be analyzed by previous methods.

5. CONCLUSION

When Sylvester's theorem is applied in unimodular 2×2 transfer-matrix optics, it governs light propagation through periodic optical systems. There are numerous physically interesting and important applications of these periodic systems. However, some applications, such as those involving periodic distributed-feedback lasers, lossy birefringent filters, electric circuits with intranetwork independent sources, high-energy particle accelerators, periodic computer graphics manipulations with object translations, and periodic pulse compressors, require nonunimodular and/or an augmented 3×3 matrix formalism. We have extended Sylvester's theorem here to facilitate the study of these and other systems. We have also used systematic procedures to find the range of validity of Sylvester's theorem, and several important special cases have been identified. For example, roots of matrices, which are useful for system synthesis, were examined. The basic results have been summarized in tabular form, and it is straightforward to extend these results to other types of matrix.

APPENDIX A: DERIVATION OF SYLVESTER'S THEOREM

The purpose of this appendix is to derive the basic unimodular Sylvester theorem from first principles with the use of only algebra and simple matrix multiplication. This treatment is a simplified version of the derivation given in Ref. 1.

Suppose that a unimodular matrix T is written as a product of three matrices in the form

$$T = M^{-1} \Lambda M. \quad (\text{A1})$$

This is a particularly attractive form, since one exactly finds that

$$T^s = M^{-1} \Lambda^s M \quad (\text{A2})$$

because of the definition of the matrix inverse. For simplicity, the matrix Λ is chosen to be diagonal, and therefore

$$\Lambda^s = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}^s = \begin{bmatrix} \lambda_1^s & 0 \\ 0 & \lambda_2^s \end{bmatrix}. \quad (\text{A3})$$

Thus the procedure is to solve for the Λ and M matrices. To begin, we multiply both sides of Eq. (A1) on the left by M to obtain

$$\begin{bmatrix} M_1 & M_2 \\ M_3 & M_4 \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} M_1 & M_2 \\ M_3 & M_4 \end{bmatrix}. \quad (\text{A4})$$

Multiplying these out yields four equations with six un-

knowns. It therefore follows that there are an infinite number of choices for the matrix M . The four equations may be reformulated as two matrix equations:

$$\begin{bmatrix} A - \lambda_1 & C \\ B & D - \lambda_1 \end{bmatrix} \begin{bmatrix} M_1 \\ M_2 \end{bmatrix} = 0, \quad (\text{A5})$$

$$\begin{bmatrix} A - \lambda_2 & C \\ B & D - \lambda_2 \end{bmatrix} \begin{bmatrix} M_3 \\ M_4 \end{bmatrix} = 0. \quad (\text{A6})$$

Equation (A5) is satisfied when M_1 and M_2 are both zero, but in this case the matrix M has a zero determinant. This violates the assumption in Eq. (A1) of the existence of M^{-1} . Since M_1 and M_2 are not both zero, the determinant of the matrix in Eq. (A5) must be zero. A similar argument may be used for Eq. (A6). The requirement that the determinants of these two matrices must be zero implies two independent and identical equations with two unknowns. These equations are sometimes referred to as eigenvalue equations, and λ_1 and λ_2 are known as eigenvalues. The eigenvalue equations are

$$\lambda_{1,2}^2 - (A + D)\lambda_{1,2} + 1 = 0, \quad (\text{A7})$$

where the fact that the matrix T is unimodular ($AD - BC = 1$) has been used. The eigenvalues are

$$\lambda_{1,2} = \exp(\pm i\theta), \quad (\text{A8})$$

where the definition

$$\cos \theta = \frac{A + D}{2} \quad (\text{A9})$$

has been used.

Now that the eigenvalues λ_1 and λ_2 have been found, the goal is to find the matrix M . The top rows of Eqs. (A5) and (A6) may be expressed as

$$M_2 = \frac{(\lambda_1 - A)M_1}{C}, \quad (\text{A10})$$

$$M_4 = \frac{(\lambda_2 - A)M_3}{C}. \quad (\text{A11})$$

However, as we mentioned above, there are two more unknowns than equations. Therefore M_1 and M_3 are arbitrarily chosen to be unity. Thus the matrix M is

$$M = \begin{bmatrix} 1 & (\lambda_1 - A)/C \\ 1 & (\lambda_2 - A)/C \end{bmatrix}, \quad (\text{A12})$$

and the inverse of M is

$$M^{-1} = \frac{C}{\lambda_2 - \lambda_1} \begin{bmatrix} (\lambda_2 - A)/C & -(\lambda_1 - A)/C \\ -1 & 1 \end{bmatrix}. \quad (\text{A13})$$

Now that the matrices λ and M have been calculated, the s th power of the matrix T is given by Eq. (A2):

$$T^s = \frac{C}{\lambda_2 - \lambda_1} \begin{bmatrix} (\lambda_2 - A)/C & -(\lambda_1 - A)/C \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1^s & 0 \\ 0 & \lambda_2^s \end{bmatrix} \begin{bmatrix} 1 & (\lambda_1 - A)/C \\ 1 & (\lambda_2 - A)/C \end{bmatrix} \quad (\text{A14})$$

$$= \frac{1}{\lambda_2 - \lambda_1} \begin{bmatrix} \lambda_1^s(\lambda_2 - A) - \lambda_2^s(\lambda_1 - A) & (\lambda_1 - A)(\lambda_2 - A)(\lambda_2^s - \lambda_1^s)/C \\ C(\lambda_2^s - \lambda_1^s) & \lambda_2^s(\lambda_2 - A) - \lambda_1^s(\lambda_1 - A) \end{bmatrix}. \quad (\text{A15})$$

Making using of the unimodularity condition ($AD - BC = 1$) and Eqs. (A8) and (A9), one may show that

$$\lambda_1 + \lambda_2 = A + D, \quad (\text{A16})$$

$$\frac{\lambda_2^s - \lambda_1^s}{\lambda_2 - \lambda_1} = \frac{\sin(s\theta)}{\sin \theta}. \quad (\text{A17})$$

We may use these relations to reduce Eq. (A15) to

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^s = \frac{1}{\sin \theta} \times \begin{bmatrix} A \sin(s\theta) - \sin[(s-1)\theta] & B \sin(s\theta) \\ C \sin(s\theta) & D \sin(s\theta) - \sin[(s-1)\theta] \end{bmatrix}, \quad (\text{A18})$$

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{s+1} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix}^s \quad (\text{B4})$$

$$= \frac{1}{\sin \theta} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} A \sin(s\theta) - \sin[(s-1)\theta] & B \sin(s\theta) \\ C \sin(s\theta) & D \sin(s\theta) - \sin[(s-1)\theta] \end{bmatrix}. \quad (\text{B5})$$

$$= \frac{1}{\sin \theta} \begin{bmatrix} (A^2 + BC)\sin(s\theta) - A \sin[(s-1)\theta] & B(A + D)\sin(s\theta) - B \sin[(s-1)\theta] \\ C(A + D)\sin(s\theta) - C \sin[(s-1)\theta] & (BC + D^2)\sin(s\theta) - D \sin[(s-1)\theta] \end{bmatrix}. \quad (\text{B6})$$

which is Sylvester's theorem. It should be noted that, in this derivation, there has been no overt restriction that s be an integer. As we mentioned in Section 2 above, after the inductive proof of Sylvester's theorem (Appendix B below), roots of matrices are specifically examined. It is found that Sylvester's theorem applies to roots of matrices (Appendix C) and rational powers of matrices (Appendix D).

APPENDIX B: PROOF OF SYLVESTER'S THEOREM

One may use mathematical induction to prove Sylvester's Theorem.³² For our purposes Sylvester's theorem can be formulated as follows:

Given a 2×2 matrix that is unimodular, i.e.,

$$AD - BC = 1, \quad (\text{B1})$$

the s th power of that matrix can be expressed as

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^s = \frac{1}{\sin \theta} \times \begin{bmatrix} A \sin(s\theta) - \sin[(s-1)\theta] & B \sin(s\theta) \\ C \sin(s\theta) & D \sin(s\theta) - \sin[(s-1)\theta] \end{bmatrix}, \quad (\text{B2})$$

where

$$\cos \theta = \frac{A + D}{2}. \quad (\text{B3})$$

The induction proof of this theorem follows these three steps:

(1) Demonstrate the validity of the theorem for the case $s = 1$.

(2) Assume the validity of the theorem for some arbitrary s th case.

(3) Demonstrate the validity of the theorem for the $(s + 1)$ th case.

The first step follows by inspection. The second step is a restatement of the theorem [Eqs. (B1)–(B3)]. For the last step the $(s + 1)$ th power of a matrix is calculated as

To continue, it is easiest to examine each of the four matrix elements individually. The first matrix element is

$$A_{s+1} = (A^2 + BC)\sin(s\theta) - A \sin[(s-1)\theta] \quad (\text{B7})$$

$$= A(A + D)\sin(s\theta) - A[\sin(s\theta)\cos \theta - \cos(s\theta)\sin \theta] - \sin(s\theta) \quad (\text{B8})$$

$$= A[\sin(s\theta)\cos \theta + \cos(s\theta)\sin \theta] - \sin(s\theta) \quad (\text{B9})$$

$$= A \sin[(s + 1)\theta] - \sin(s\theta), \quad (\text{B10})$$

where the definition of θ [Eq. (B3)] has been used. The second matrix element is

$$B_{s+1} = B(A + D)\sin(s\theta) - B \sin[(s-1)\theta] \quad (\text{B11})$$

$$= 2B \sin(s\theta)\cos \theta - B[\sin(s\theta)\cos \theta - \cos(s\theta)\sin \theta] \quad (\text{B12})$$

$$= B[\sin(s\theta)\cos \theta + \cos(s\theta)\sin \theta] \quad (\text{B13})$$

$$= B \sin[(s + 1)\theta], \quad (\text{B14})$$

where again the definition of θ [Eq. (B3)] has been used. The third matrix element, C_{s+1} , is identical to B_{s+1} if B and C are interchanged. Similarly, the fourth matrix element, D_{s+1} , is identical to A_{s+1} if A and D are interchanged. Thus the $(s + 1)$ th power of a unimodular 2×2 matrix is

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{s+1} = \frac{1}{\sin \theta} \times \begin{bmatrix} A \sin[(s+1)\theta] - \sin(s\theta) & B \sin[(s+1)\theta] \\ C \sin[(s+1)\theta] & D \sin[(s+1)\theta] - \sin(s\theta) \end{bmatrix}. \quad (\text{B15})$$

But this is Eq. (B2) with s replaced by $s+1$. Thus Sylvester's theorem has been proved for $s=1$ and every integer greater than unity. The steps in the proof may be performed in reverse order, and it follows that Sylvester's theorem is valid for $s=0$ and negative integers as well. In the cases $s=0$ and $s=1$ the matrix T^s in Sylvester's theorem reduces to the identity matrix and the inverse matrix, respectively. It may also be proved, as is done in Appendix C below, that Sylvester's theorem is valid for nonintegers as well.

APPENDIX C: ROOTS OF MATRICES

Since it has been established that Sylvester's theorem

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A_s & B_s \\ C_s & D_s \end{bmatrix}^{1/s} = \frac{1}{\sin \phi} \begin{bmatrix} A_s \sin(\phi/s) - \sin[(1/s-1)\phi] & B_s \sin(\phi/s) \\ C_s \sin(\phi/s) & D_s \sin(\phi/s) - \sin[(1/s-1)\phi] \end{bmatrix}, \quad (\text{C9})$$

applies to integer powers of a matrix, the purpose of this appendix is to show that Sylvester's theorem applies to roots of matrices as well.

Sylvester's theorem is

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^s = \begin{bmatrix} A_s & B_s \\ C_s & D_s \end{bmatrix} = \frac{1}{\sin \theta} \times \begin{bmatrix} A \sin(s\theta) - \sin[(s-1)\theta] & B \sin(s\theta) \\ C \sin(s\theta) & D \sin(s\theta) - \sin[(s-1)\theta] \end{bmatrix}, \quad (\text{C1})$$

where

$$\cos \theta = \frac{A+D}{2}. \quad (\text{C2})$$

We may use these two equations to obtain the integer power of a matrix. In this case A , B , C , and D are assumed to be known, and the problem consists in solving for A_s , B_s , C_s , and D_s . However, if A_s , B_s , C_s , and D_s were known, then solving for A , B , C , and D would be equivalent to finding the s th root of a matrix. Thus the problem is to invert Sylvester's theorem [Eqs. (C1) and (C2)]. Equation (C1) may be rewritten as

$$(\sin \theta) \begin{bmatrix} A_s & B_s \\ C_s & D_s \end{bmatrix} = \sin(s\theta) \begin{bmatrix} A & B \\ C & D \end{bmatrix} - \sin[(s-1)\theta] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (\text{C3})$$

When we rearrange Eq. (C3), it follows that

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \frac{1}{\sin(s\theta)} \times \begin{bmatrix} A_s \sin \theta + \sin[(s-1)\theta] & B_s \sin \theta \\ C_s \sin \theta & D_s \sin \theta + \sin[(s-1)\theta] \end{bmatrix}. \quad (\text{C4})$$

Also, from Eq. (C1) it can be seen that

$$\frac{A_s + D_s}{2} = \frac{[(A+D)/2]\sin(s\theta) - \sin[(s-1)\theta]}{\sin \theta} \quad (\text{C5})$$

$$= \frac{(\cos \theta)\sin(s\theta) - \sin[(s-1)\theta]}{\sin \theta} \quad (\text{C6})$$

$$= \cos(s\theta), \quad (\text{C7})$$

where Eq. (C2) has been used. If the substitution

$$\phi = s\theta \quad (\text{C8})$$

is made, then Eq. (C4) becomes

and Eq. (C7) reduces to

$$\cos \phi = \frac{A_s + D_s}{2}. \quad (\text{C10})$$

But Eqs. (C9) and (C10) are identical in form to Sylvester's theorem [Eqs. (C1) and (C2)] when the matrix power s is replaced with $1/s$. It therefore follows that Sylvester's theorem is valid for roots of matrices as well as integer powers of matrices. That it makes sense to speak of matrix roots for nonzero-determinant matrices was recognized by Cayley.²⁵ In Eq. (C10), $\cos \phi = \cos(-\phi) = \cos(2k\pi - \phi)$, and there is more than one possible value for $\sin(\phi/s)$. This may be accounted for by multiplication of the right-hand side of Eq. (C9) by $1^{1/s}$.

APPENDIX D: RATIONAL POWERS OF MATRICES

Since it has been established that Sylvester's theorem applies to integer powers and roots of a matrix, the purpose of this appendix is to show that Sylvester's theorem also applies to rational powers of matrices. Since one may use rational numbers to approximate an irrational number with arbitrary precision, it will follow that Sylvester's theorem also applies to arbitrary powers of a matrix.

If it could be shown that

$$(T^n)^m = T^{nm}, \quad (\text{D1})$$

then it would follow that

$$(T^n)^{1/s} = T^{n/s}. \quad (\text{D2})$$

Since Sylvester's theorem has been proved to apply to

both integer powers and roots of matrices, it would follow from Eq. (D2) that it applies to rational powers of a matrix as well.

Thus the purpose here is to prove Eq. (D1). According to Sylvester's theorem,

$$T^n = \begin{bmatrix} A_n & B_n \\ C_n & D_n \end{bmatrix} = \frac{1}{\sin \theta} \times \begin{bmatrix} A \sin(n\theta) - \sin[(n-1)\theta] & B \sin(n\theta) \\ C \sin(n\theta) & D \sin(n\theta) - \sin[(n-1)\theta] \end{bmatrix}, \quad (\text{D3})$$

where

$$\cos \theta \equiv \frac{A + D}{2}. \quad (\text{D4})$$

Therefore

$$(T^n)^m = \begin{bmatrix} A_{nm} & B_{nm} \\ C_{nm} & D_{nm} \end{bmatrix} = \frac{1}{\sin \phi} \begin{bmatrix} A_n \sin(m\phi) - \sin[(m-1)\phi] & B_n \sin(m\phi) \\ C_n \sin(m\phi) & D_n \sin(m\phi) - \sin[(m-1)\phi] \end{bmatrix}, \quad (\text{D5})$$

where

These results may be combined as

$$(T^n)^m = \frac{1}{\sin \theta} \begin{bmatrix} A \sin(nm\theta) - \sin[(nm-1)\theta] & B \sin(nm\theta) \\ C \sin(nm\theta) & D \sin(nm\theta) - \sin[(nm-1)\theta] \end{bmatrix}, \quad (\text{D18})$$

$$\cos \phi \equiv \frac{A_n + D_n}{2}. \quad (\text{D6})$$

where

$$\cos \theta \equiv \frac{A + D}{2}. \quad (\text{D19})$$

The angles θ and ϕ are related through Eqs. (D3), (D4), and (D6). Equation (D6) reduces to

$$\cos \phi = \frac{[(A + D)/2] \sin(n\theta) - \sin[(n-1)\theta]}{\sin \theta} \quad (\text{D7})$$

$$= \frac{(\cos \theta) \sin(n\theta) - \sin[(n-1)\theta]}{\sin \theta} \quad (\text{D8})$$

$$= \cos(n\theta). \quad (\text{D9})$$

It therefore follows that

$$\phi = n\theta. \quad (\text{D10})$$

It is easiest to proceed in the reduction of Eq. (D5) if we look at one matrix element at a time. The first matrix element is

$$A_{nm} = \frac{A_n \sin(m\phi) - \sin[(m-1)\phi]}{\sin \phi} \quad (\text{D11})$$

$$= \frac{\left\{ \frac{A \sin(n\theta) - \sin[(n-1)\theta]}{\sin \theta} \right\} \sin(nm\theta) - \sin[(m-1)n\theta]}{\sin(n\theta)} \quad (\text{D12})$$

$$= \frac{A \sin(nm\theta)}{\sin \theta} - \frac{\sin[(nm-1)\theta]}{\sin \theta} - \frac{\sin[(n-1)\theta] \sin(nm\theta) + (\sin \theta) \sin[(m-1)n\theta] - \sin(n\theta) \sin[(nm-1)\theta]}{\sin(n\theta) \sin \theta} \quad (\text{D13})$$

$$= \frac{A \sin(nm\theta)}{\sin \theta} - \frac{\sin[(nm-1)\theta]}{\sin \theta}, \quad (\text{D14})$$

where Eq. (D10) has been used together with a basic trigonometric identity. For simplicity, the second term in Eq. (D13) has been both added to and subtracted from the equation. The second matrix element can be found in a like manner. From Eqs. (D5) and (D10) it follows that

$$B_{nm} = \frac{B_n \sin(m\phi)}{\sin \phi} \quad (\text{D15})$$

$$= \frac{B \sin(n\theta)}{\sin \theta} \frac{\sin(nm\theta)}{\sin(n\theta)} \quad (\text{D16})$$

$$= \frac{B \sin(nm\theta)}{\sin \theta}. \quad (\text{D17})$$

The derivation of the third matrix element, C_{nm} , is identical to the derivation for B_{nm} if B is replaced with C . Similarly, the derivation of the fourth matrix element, D_{nm} , is identical to the derivation for A_{nm} if A is replaced with D .

ACKNOWLEDGMENTS

The authors acknowledge valuable discussions with Bruce D. Ulrich. This study was supported in part by the National Science Foundation under grant ECS-9014481.

REFERENCES

1. A. Gerrard and J. M. Burch, *Introduction to Matrix Methods in Optics* (Wiley, New York, 1975).

2. A. A. Tovar and L. W. Casperson, "Generalized reverse theorems for multipass applications in matrix optics," *J. Opt. Soc. Am. A* **11**, 2633–2642 (1994).
3. J. A. Arnaud, "Degenerate optical cavities. II. Effects of misalignment," *Appl. Opt.* **8**, 1909–1917 (1969).
4. J. R. Pierce, "Modes in sequences of lenses," *Proc. Natl. Acad. Sci. USA* **47**, 1803–1813 (1961).
5. L. W. Casperson and S. D. Lunnam, "Gaussian modes in high loss laser resonators," *Appl. Opt.* **14**, 1193–1199 (1975).
6. J. U. White, "Long optical paths of large aperture," *J. Opt. Soc. Am.* **32**, 285–288 (1942).
7. L. W. Casperson and P. M. Scheinert, "Multipass resonators for annular gain lasers," *Opt. Quantum. Electron.* **13**, 193–199 (1981).
8. L. W. Casperson, "Beam propagation in periodic quadratic-index waveguides," *Appl. Opt.* **24**, 4395–4403 (1985).
9. D. A. Edwards and H. J. Syphers, *An Introduction to the Physics of High Energy Accelerators* (Wiley, New York, 1993), pp. 60–68.
10. P. J. Bryant and K. Johnson, *The Principles of Circular Accelerators and Storage Rings* (Cambridge U. Press, Cambridge, 1993), p. 37.
11. A. Yariv and P. Yeh, *Optical Waves in Crystals* (Wiley, New York, 1984).
12. R. C. Jones, "A new calculus for the treatment of optical systems. I. Description and discussion of the calculus," *J. Opt. Soc. Am.* **31**, 488–493 (1941).
13. P. Yeh, "Extended Jones matrix method," *J. Opt. Soc. Am.* **72**, 507–513 (1982).
14. M. Born and E. Wolf, *Principles of Optics*, 6th ed. (Pergamon, New York, 1980), pp. 51–70.
15. J. Hong, W. Huang, and T. Makino, "On the transfer matrix method for distributed-feedback waveguide devices," *J. Lightwave Technol.* **10**, 1860–1868 (1992).
16. J. Capmany and M. A. Muriel, "A new transfer matrix formalism for the analysis of fiber ring resonators: compound coupled structures for FDMA demultiplexing," *J. Lightwave Technol.* **8**, 1904–1919 (1990).
17. O. E. Martinez, "Matrix formalism for pulse compressors," *IEEE J. Quantum Electron.* **24**, 2530–2536 (1988).
18. O. E. Martinez, "Matrix formalism for dispersive laser cavities," *IEEE J. Quantum Electron.* **25**, 296–300 (1989).
19. D. A. Edwards and H. J. Syphers, *An Introduction to the Physics of High Energy Accelerators* (Wiley, New York, 1993), pp. 88–94.
20. R. A. Plastock and G. Kalley, *Theory and Problems of Computer Graphics*, Schaum's Outline Series (McGraw-Hill, New York, 1986), pp. 82–87.
21. J. J. Sylvester, "Sur les puissances et les racines de substitutions lineaires," *C. R. Acad. Sci.* **XCIV**, 55–59 (1882); also published in *The Collected Mathematical Papers of James Joseph Sylvester* (Cambridge U. Press, Cambridge, 1912), Vol. 4, pp. 562–564.
22. P. I. Richards, *Manual of Mathematical Physics* (Pergamon, New York, 1959), pp. 311–312.
23. R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis* (Cambridge U. Press, Cambridge, 1991), p. 401.
24. S. Barnett, *Matrices, Methods and Applications* (Clarendon, Oxford, 1990), p. 234.
25. A. Cayley, "A memoir on the theory of matrices," *Philos. Trans. R. Soc. London* **CXLVIII**, 17–27 (1858); also published in *The Collected Mathematical Papers of Arthur Cayley* (Cambridge U. Press, Cambridge, 1889), Vol. 2, pp. 475–496.
26. Ref. 14, p. 67.
27. A. Yariv, *Quantum Electronics*, 3rd ed. (Wiley, New York, 1989), p. 124.
28. H. A. Haus, *Waves and Fields in Optoelectronics* (Prentice-Hall, Englewood Cliffs, N.J., 1984), pp. 138–139.
29. P. Yeh, A. Yariv, and C.-S. Hong, "Electromagnetic propagation in periodic stratified media. I. General theory," *J. Opt. Soc. Am.* **67**, 423–438 (1977).
30. H. Kogelnik, "Imaging of optical modes—resonators with internal lenses," *Bell Syst. Tech. J.* **44**, 455–494 (1965).
31. H. Kogelnik and T. Li, "Laser beams and resonators," *Proc. IEEE* **54**, 1312–1329 (1966).
32. P. W. Milonni and J. H. Eberly, *Lasers* (Wiley, New York, 1988), pp. 476–477.
33. R. A. Fraser, W. J. Duncan, and A. R. Collar, *Elementary Matrices and Some Applications to Dynamics and Differential Equations* (Cambridge U. Press, Cambridge, 1963), pp. 83–87.
34. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1964), pp. 771–802; see also M. R. Spiegel, *Mathematical Handbook of Formulas and Tables*, Schaum's Outline Series (McGraw-Hill, New York, 1968), pp. 157–159.
35. L. W. Casperson, "Gaussian light beams in inhomogeneous media," *Appl. Opt.* **12**, 2434–2441 (1973).
36. J. T. Verdeyen, *Laser Electronics*, 2nd ed. (Prentice-Hall, Englewood Cliffs, N.J., 1989), pp. 38–48.